Modified node identification using Convex Hull & K-Mean approach in Mobile Ad Hoc Networks

A Synopsis

Submitted in

Partial fulfillment for the degree

Of

Doctor of Philosophy

(Computer Science & Engineering)

Supervised by

JV'n Dr. Kavita

Submitted by

JV'n Shilpa Pareek

JVR-1/21/5822

Department of Computer Science and Information Technology

Faculty of Engineering & Technology

Jayoti Vidyapeeth Women's University

Jaipur (Rajasthan), India

Introduction

In MANET (mobile ad-hoc network) nodes are self-organized and they are communicating between each other without any stable infrastructure. Now when in MANET network, the path chosen for the routing protocol faces malicious attack many times. If the path is created for n ton communication this path is also not stable because nodes are basically mobile which always change the position. Now here we describe a new on-demand secure routing protocol using data mining approach where if a node is suddenly missing or the position is changed it will not hamper the routing path. This protocol is for find shortest path and checks how many least nodes are vital for a n to n communication. According to that, firstly the cluster is created using K-mean. Now those nodes in the cluster, using convex hull, identify the outer region of this cluster. From this average of the outer region cluster the proper node for the routing path is chosen and the same procedure is continued for each and every node selection in the on-demand routing path.

A wireless ad-hoc network, also known as IBSS - Independent Basic Service Set, is a computer network in which the communication links are wireless. The network is ad-hoc because each node wants to forward data for other nodes, and so the determination of which nodes forward data is made dynamically based on the network connectivity. This is in contrast to older network technologies in which some designated nodes, usually with custom hardware and variously known as routers, switches, hubs, and firewalls, perform the task of forwarding the data. Minimal configuration and quick deployment make ad hoc networks suitable for emergency situations like natural or human-induced disasters, military conflicts.

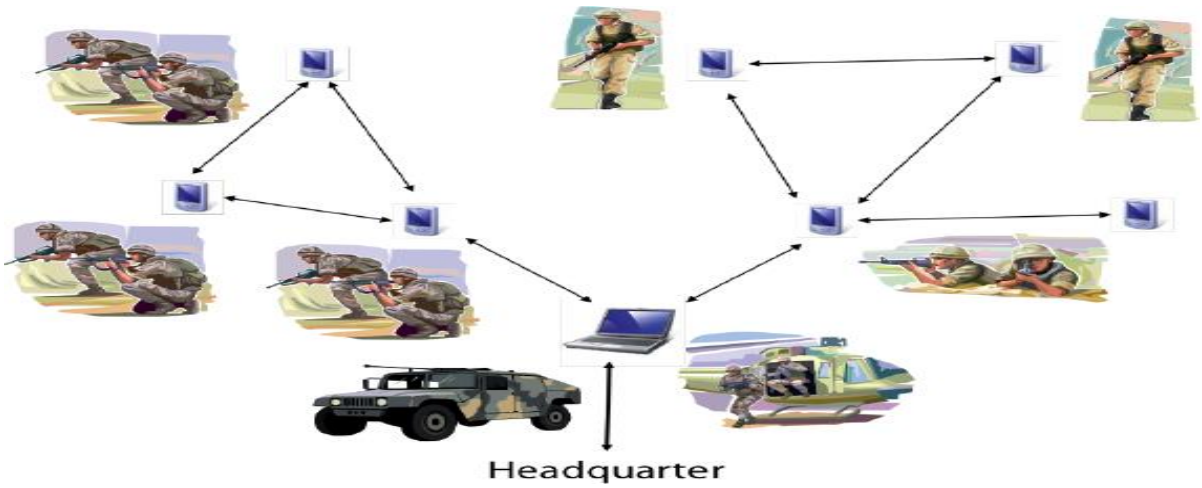

Fig1. Overview of Mobile Ad-hoc Network

Brief Literature Survey:

A.Sharma *et al* ((2018) Have present a novel routing algorithm for wireless sensor network, which achieves uniform energy depletion across all the nodes and thus leading to prolonged network lifetime. The proposed algorithm, divides the Region of Interest into virtual zones, each having some designated cluster head nodes. In the entire process, a node can either be a part of a cluster or it may remain as an independent entity. A non-cluster member transmits its data to next hop node using IRP-Intelligent Routing Process (based on the trade-off between the residual energy of itself as well as its neighbor, and the required energy to transmit packets to its neighbor). If on the transmission path, some cluster member is elected as a next hop, it rejects IRP and transmits the packets to cluster head, which later forwards them to sink (adopting multihop communication among cluster heads). Routing is not solely performed using clusters, rather they aid the overall routing process, hence this protocol is named as Cluster Aided Multipath Routing (CAMP). CAMP has been compared with various sensor network routing protocols, viz., LEACH, PEGASIS, DIRECT TRANSMISSION, CEED, and CBMR. It is found that the proposed algorithm outperformed them in network lifetime, energy consumption and coverage ratio.

R. Dey *et al.* (2015) Have present structured data that is efficiently and effectively used by the DBSCAN clustering algorithm. Machine learning methods are increasingly being used in conjunction with conventional meteorological observations in the synoptic analysis and conventional weather forecast to extract information of relevance for agriculture and food security of the human society in India. Density based clustering approach is incrementally used to predict the future weather conditions for this project. One famous preprocessing approach, known as Convex-Hull is also used before fed the pollutant data into the clustering algorithm. This Convex-Hull method is strictly used to convert unstructured data into its corresponding structured form. These structured data is efficiently and effectively used by the DBSCAN clustering algorithm to form resultant clusters for weather derivatives. This forecasting database is totally based on the weather of Kolkata city in west Bengal and this forecasting methodology is developed to mitigating the impacts of air pollutions and launch focused modeling computations for prediction and forecasts of weather events. Here accuracy of this approach is also measured.

S. Banerjee *et al*. (2015) Have present a protocol is based on swarm intelligence and especially on the ant colony based Meta heuristic. A mobile ad-hoc network (MANET) is a collection of mobile nodes which communicate over radio. These kinds of networks are very flexible, thus they do not require any existing infrastructure or central administration. Therefore, mobile ad-hoc networks are suitable for temporary communication links. The biggest challenge in this kind of networks is to find a path between the communication end points, which is aggravated through the node mobility. In this paper we present a new on-demand power-balanced routing algorithm

for mobile, multi-hop ad-hoc networks. The protocol is based on swarm intelligence and especially on the ant colony based meta heuristic. These approaches try to map the solution capability of swarms to mathematical and engineering problems. The proposed routing protocol is highly adaptive, efficient and scalable. The main goal in the design of the protocol is to reduce the overhead for routing. Our simulation results show that the proposed routing protocol is significantly different from existing protocols.

R.Dey *et al.* (2015) Have discuss about MANET and its vulnerabilities .In recent years, Mobile Ad hoc Network (MANET) have become a very popular research topic. By providing communications in the absence of a fixed infrastructure MANET are an attractive technology for many applications such as resource app, military app, environmental monitoring and conferences. However, this flexibility introduces new security threats due to the vulnerable nature of MANET, there will be the necessity of protecting the data, information from the attackers as it is an infrastructure-less network. Thus, securing such demanding network is a big challenge. At this point, IDS came into existence to secure MANET in detecting at what point they are getting weak. In this review paper, we will discuss, MANET and its vulnerabilities, and how we can tackle it using different techniques of IDS (Intrusion Detection System).

R.Dey *et al.* (2014) Have present a a new technique to predict the weather conditions using convex hull. Data mining is a popular concept of mined necessary data from a large set of data. Data mining using clustering is a powerful way to analyze data and gives prediction. This is non-structural time series data is used to forecast daily average temperature, humidity and overall weather conditions of Kolkata city. The air pollution data have been taken from West Bengal Pollution Control Board to build the original dataset on which the prediction approach of this paper is studied and applied. This paper describes a new technique to predict the weather conditions using convex hull which gives structural data and then apply incremental K-means to define the appropriate clusters. It splits the total database into four separate databases with respect to different weather conditions. In the final step, the result will be calculated on the basis of priority based protocol which is defined based on some mathematical deduction.

Elizabeth *et al.* (2012) (SRT-Secure Routing Using Trust) have presented a protocol which uses trust level which is based on node transition probability (NTP) and AODV. This scheme tries to secure NTP protocol. A trust rate (Trate) is calculated as a parameter. When a node has data packet to send, it first floods control frame (beacon) in search of a secure and reliable route, after which the trust rate is evaluated. This Trate value divides the nodes of the network into 3 categories: Ally list (Level 2), which send highly secured information; Associate list (Level 1),which send medium secured information; and Acquaintance list(Level 0), which send the information that do not require any security. The advantage is that it is able to provide high throughput and PDF, but the disadvantage is that the performance decreases in the presence of malicious activities, except black hole attack. Moreover the trust is calculated only on the basis of control packets, and the battery power is not considered.

Saha *et al.* (2010) have proposed Semi-Centralized Multi Authenticated RSSI Based Solution to Sybil Attack. It provides a hybrid solution where the entire network is subdivided into small groups under the supervision of each nodes. Here each of the subgroups have RSSI detector nodes. The advantage is that even when false reading is sent, the problem will be restricted in the group only. The demerit is that if any of detectors in the group can be compromised then there will be trouble, also the scheme does not address any energy constraints and can't tackle other attacks other than Sybil.

Saha *et al.* (2010) have proposed Priority Based Protocol which has capacity to detect malicious nodes with no notification to central authority. Here the nodes will maintain priority list of neighbors and when the value will go below the threshold level, the node will be disconnected from the system. Nodes here can also upgrade themselves and act as per node, as they are independent of CA. The merit is that this protocol can able to detect different attacks like Black hole attack, Gray attack, but fails to combat Jellyfish. This scheme can be used in distributed manner which provides less network dependence of nodes.

Patwardhan and **Yorga *et al*. (2005)** have proposed the secured routing protocol model called Sec AODV, which is modified form of AODV, where each nodes are provided with a certain IPv6 address value within the MANET. In this scheme, the secured communication channel is established in between source and destination node depending upon the idea of Statistically Unique and Cryptographically Verifiable (SUCV) who confirms secured and secret binding between IPv6 address and key node. [54] In this scheme, there is no requirement of trust certificate (CA). Still this protocol is restricted in working zone since here it is necessary for all the MANET nodes to contain a certain IPV6 address. It fails to stop or prevent the selective dropping of packets, as the probing packets can be easily recognized from others. Along with this, scheme does not have address for energy constrains.

Nekkauti and **Lee *et al.* (2004)** have given birth to a certain trust based adaptive routing protocol. This protocol uses the mechanism of conversion of the codes into order while covering the route path between the source and destination through all other nodes. This scheme supports greater degree of individualism of nodes to operate in routing path but the greatest disadvantage is it can't protect network system from malicious activities where there is the selective dropping of packets which are agreed to be forwarded, under malicious nodes.

Yi *et al.* (2002) have proposed security aware adhoc routing also termed as SAR. In his proposition, SAR has classified the nodes on basis of their trust values. On the basis of this classification, the similar nodes with same trust level share a secret group key. In determining route, the source node agrees to have minimum number of security requirements, which a node generally requires in order to complete its routing path from source to destination. Source node can also agree to make path by encryption of route request packets with certain specified keys for specific levels. In it keys become the main defect for the scheme, which can create problems

when malicious nodes will control the keys in spite of having high security while controlling the secret group keys.

Objectives:

The objectives associated with the work are listed below:

- To develop an algorithm to retain the energy of node in remote area (military people) for future use.
- To construct an on demand routing path protocol that will be loop free, lightweight reliable, robust, energy efficient and uses less memory space.
- To create a suitable dynamic databases where the positions of node are changed frequently.
- To make a probable guess of the initial cluster center to get more effective results.

Significance of Work:

The significance of the proposed model for on demand routing protocol on the MANET provided. The features of this routing protocol are given here under:

- In MANET, most of the nodes are depending upon battery system, so their energy is limited. But in this proposed protocol it uses priority like high threshold value, medium threshold value, and low threshold value cluster head.
- The method will be able to easily identify the good node to transfer the information in remote area.
- The On demand routing path protocol will be robust, energy efficient and uses less memory space as well as loop free, Light weight and reliable .
- The k-means algorithm to identify the cluster and the cluster center should not be malicious node. Cluster head will decide the trust value of the entire cluster members present in the cluster.
- The Convex Hull is ensures to take all different extreme values in the database. This is basically graph based calculation to find out the basic structure of the database.

Methodology:

In this protocol we collect the entire nodes within the range of the source towards the destination. Each and every node has certain capacity to send data within a certain distance. Now data reach from source to next node towards the destination. For choosing first node to next node, how many available nodes are calculated, and then according to availability use k means algorithm to identify the cluster and the cluster center should not be malicious node. Cluster head will decide the trust value of the entire cluster members present in the cluster.

Then use of convex hull (upper hull and lower hull) identifies the outer region of this cluster. From this outer region node, it takes the average value of those nodes and the nearest node of this average value is chosen, which is used for data transfer. This protocol is repeated again and again until it reaches the destination node. From the next time of data transfer, previous chosen node is treated as highest priority node so that next time, selection takes less time.

Cluster: Cluster analysis divides data into groups in the basis of meaningful and useful. Cluster analysis assemblages data objects based only on information initiate in the data that describes the objects and their associations or relationship with each other. The goal is that the objects within a group be similar, related to one another and deferent from the objects in other groups. The more the similarity within a group and the more the difference between groups, the better or more distinct the clustering.

K-means Clustering: The K-means clustering technique is simple. This algorithm first chooses N initial centroids, where N is a user-specified framework, namely, the number of clusters preferred. Every node is then assigned to the neighboring node i.e. called centroid, and each collection of nodes allotted to a centroid is a group i.e. called cluster. The centroid of each cluster is then updated based on the points assigned to the cluster. The centroids remain the same until and unless we modified the assignment and update steps and no point changes clusters, or equivalently.

Convex Hull: This procedure ensures to take all different extreme values in the database. Basically this is a graph based calculation to find out the required structure of the database. It is the smallest convex set of some given points. Firstly the point of the plot source to destination on x, y plane graph. Choose right and left extreme point (P_R, P_L). P_0 is the starting point and it goes towards the immediate next point. If the next point goes towards the same direction or goes to right angle on the basis of previous direction then it will be accepted. If next point goes towards the left angle on the basis of previous direction, it rejects this node and previous node, and chooses different nodes, complete upper part of the graph. The same case is made in the lower part Figure 2. In large data analysis, it gives the fast middling idea including all paramount data points.

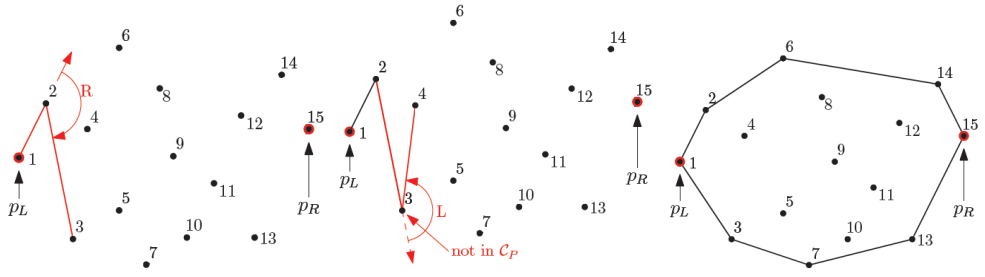

Fig.2: Convex Hull with Left angle and right angle.

Step 1: Identify all the nodes from source to destination.
Step 2: Use K-Means algorithm to identify the cluster.
Step 3: Use Convex Hull to identify outer region of the cluster.
Step 4: Identify the proper node to send data.
Step 5: Modified the priority of the selected node.
Step 6: Identify proper node and connect source to destination

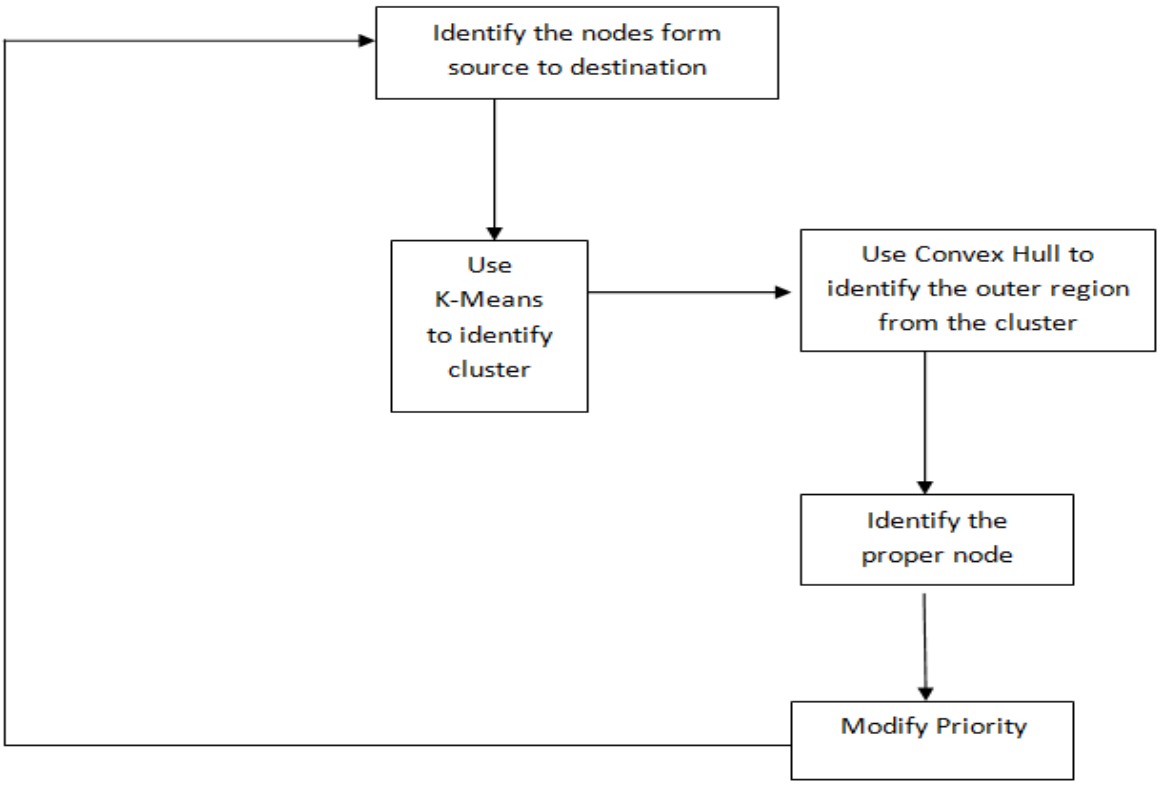
Identify the nodes form source to destination
Use K-Means to identify cluster
Use Convex Hull to identify the outer region from the cluster
Identify the proper node
Modify Priority

Fig.3: Proposed Protocol

In this protocol we collect the entire nodes within the range of the source towards the destination. Each and every node have certain capacity to send data within a certain distance. Now data reach from source to next node towards the destination. For choosing first node to next node, how many available nodes are calculated, and then according to availability use k-means algorithm to identify the cluster and the cluster center should not be malicious node. Cluster head will decide the trust value of the entire cluster members present in the cluster. Then use of convex hull (upper hull and lower hull) identifies the outer region of this cluster. From this outer region node, it takes the average value of those nodes and the nearest node of this average value is chosen, which is used for data transfer. This protocol is repeated again and again until it reaches the destination node. From the next time of data transfer, previous chosen node is treated as highest priority node so that next time, selection takes less time.

Here all the nodes are divided into three parts according to the threshold value. If the threshold value is high, it means that node is used for data transfer previously, so it is trusted node. If the threshold value is medium, it means the node is trusted but the traffic rate is high. If the node value is low, it shows the node is not used previously so it can't predict. If one low threshold value node successfully transfers the data, it is treated as high threshold value node, and also if high threshold value node drops the packet it is treated as low threshold value node. If good nodes are which transferred the data previously are missing, then cluster head choose nearest another node from the cluster.

Main references:

1) A. Yasinsac and S. Carter, "Secure Position Aided Ad hoc Routing," Florida State University, 2003.
2) H. N. Saha, D. Bhattacharyya, P. K. Banerjee, "Semi-Centralized Multi-Authenticated RSSI Based Solution to Sybil Attack", International Journal of Computer Science & Emerging Technologies (E-ISSN:2044-6004) Volume I, Issue 4, December 2010, pp. 338-341.
3) H. N. Saha, D. Bhattacharyya, P. K. Banerjee," A Priority Based Protocol for Mitigating Different Attacks in Mobile Ad hoc Networks", International Journal of Computer Science & Communication Vol. 1, No. 2, December 2010, pp. 299-302.
4) Himadri Nath Saha, Dr. Debika Bhattacharyya Dr. P.K. Banerjee, "Fidelity Index Based on Demand Secure (FBOD) Routing in Mobile Adhoc Network." In International Journal of advanced computer science and Applications.
5) K. Sanzgiri, B. Dahill, B. N. Levine, C. Shields and E. M. Belding-Royer, "A secure routing protocol for ad hoc networks,"*in proc. of 10th IEEE International Conference on Network Protocols (ICNP),* November 2002, pp.78-87.
6) L. Venkatraman and D. P. Agrawal, "An optimized inter-router authentication scheme for ad hoc networks,"*in proc. of the Wireless, 2001*, pp.129–146.
7) Mohit Sajwan, Devashish Gosain & Ajay K. Sharma " CAMP: cluster aided multi-path routing protocol for wireless sensor networks"2018
8) R. Dey, S.Chakraborty, "Convex-hull & DBSCAN clustering to predict future weather" in proc. of 6th International Conference and Workshop on Computing and Communication (IEMCON -2015), IEEE Xplore Digital Library, October 2015, pp. 1-8.
9) R. Dey, H.N. Saha, "Secure Routing Protocols for Mobile Ad-Hoc Network (MANETs) –A Review", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS) Vol.5, Issue 1, pp.74-79, 2016.
10) R.Dey, H.N.Saha, "Different Routing Threats and its Mitigations Schemes for Mobile ad-hoc Networks (MANETs) –A Review", IPASJ International Journal of Electronics & Communication (IIJEC), Vol4, Issue 3,pp. 27-38, 2016.
11) S. Dabideen and J. J. Garcia-Luna-Aceves, "Secure routing in MANET using local times," *Wireless Network,Springer*, pp.811-826, 2012.

12) S.Choi, D.Y.Kim, D.Y.Lee, and J.I.Jung “WAP: Attack Prevention Algorithm in Mobile Ad Hoc Network” IEEE International Conference on Sensor Network, Ubiquitous and Trustworthy Computing, June 2008, pp. 343-348.